

## Comparative analysis of cultivated melon groups (*Cucumis melo* L.) using random amplified polymorphic DNA and simple sequence repeat markers\*

Jack E. Staub<sup>1,\*</sup>, Yael Danin-Poleg<sup>2</sup>, Gennaro Fazio<sup>1</sup>, Thomas Horejsi<sup>1</sup>, Noa Reis<sup>2</sup> & Nurit Katzir<sup>2</sup>

<sup>1</sup>USDA-Agricultural Research Service, Vegetable Crops Unit, Department of Horticulture, 1575 Linden Dr., University of Wisconsin, Madison, WI 53706 U.S.A.; <sup>2</sup>Newe Ya'ar Research Center, Agricultural Research Organization, Ramat Yishay, 30095 Israel; (\*author for correspondence)

Received 4 May 1999; accepted 14 February 2000

**Key words:** genetic distance, germplasm management, multi-dimensional scaling, plant germplasm, RAPD, SSR

### Summary

Random amplified polymorphic DNA (RAPD) and simple sequence repeat (SSR) markers were used to characterize genetic relationships among 46 accessions in two *C. melo* L. subsp. *melo* (Cantalupensis, Inodorus) and subsp. *agrestis* (Conomon, and Flexuosus) groups. Genetic distance (GD) estimates were made among and between accessions in four melon market classes [Galia, Ogen, Charentais, and Shipper (European and U.S. types)] of Cantalupensis, one market class of Inodorus (Cassaba and Honey Dew), one accession of Conomon, and one accession of Flexuosus by employing three GD estimators; simple matching coefficient, Jaccard's coefficient, and Nei's distance-D. Differences detected among 135 RAPD bands and 54 SSR bands (products of 17 SSR primers) were used to calculate GD. Band polymorphisms observed with 21 RAPD primers and 7 SSR primers were important ( $p = 0.01$ ) in the detection of genetic differences. Estimators of GD were highly correlated ( $p > 0.0001$ ;  $r_s = 0.64$  to  $0.99$ ) when comparisons were made between estimation methods within a particular marker system. Lower correlations ( $r_s = 0.17$  to  $0.40$ ) were detected ( $P > 0.001$ ) between marker systems using any one estimator. The GD of the Conomon and Flexuosus accessions was significantly different ( $p > 0.001$ ) from the mean GD of all the market classes examined. The mean GD (Jaccard's coefficient) among accessions of Ogen, Galia, Cassaba, Charentais, European shipper, and U.S. shipper groups was  $0.11 \pm 0.04$ ,  $0.33 \pm 0.09$ ,  $0.21 \pm 0.04$ ,  $0.26 \pm 0.10$ ,  $0.17 \pm 0.05$ , and  $0.22 \pm 0.08$ , respectively. Market classes were distinct ( $p > 0.001$ ), such that GDs between Galia and other accessions were the largest (mean GD 0.34 to 0.35), and GDs between Ogen and other accessions were the smallest (mean GD 0.29 to 0.30). Contrasts between the U.S. shipper cultivar Top Mark and accessions within any market class was relatively large (mean GD =  $0.42 \pm 0.06$ ). Empirical estimations of variances associated with each marker type in the accessions examined indicated that, per band, lower coefficients of variation can be attained in the estimation of GD when using RAPDs compared to SSRs. Nevertheless, the genetic relationships identified using these markers were generally similar. The disparity between the analyses of the two markers made may be related to the amount of genome coverage which is characteristic of a particular marker system and/or its efficiency in sampling variation in a population. Results of RAPD marker analysis suggest that 80 marker bands were adequate for assessing the genetic variation present in the accessions examined.

\* The cost of publishing was defrayed in part by the payment of page charges. Under postal regulations, this paper therefore must be hereby marked as *advertisement* solely to indicate this fact. Mention of a trade name, proprietary product, or specific equipment does not

constitute that a guarantee or warranty by the USDA and does not imply its approval to the exclusion of other products that may be suitable.

## Introduction

Cultivated melon (*Cucumis melo* L.) is a member of the genus *Cucumis* in the family Cucurbitaceae. Based on differences in leaf, vine, plant and fruit characters, the *C. melo* subsp. *melo* ( $x = n = 12$ ) has been further subdivided into seven horticulturally important melon groups by Whitaker & Davis (1962) after Naudin (1859). More recently Munger & Robinson (1991) have reclassified the seven horticultural groups. Two of these groups, Cantalupensis (Muskmelon) and Inodorus (Cassaba and Honeydew), are of commercial interest in the United States, and in many European, Mediterranean and Asiatic countries. Phenotypic differences have allowed melon breeders to develop distinct horticultural types specific for geographic area (e.g., adaptation, and cultural and culinary attributes) and market characteristics (e.g., shipping or fresh market). Thus, within the horticultural groups there may be more than one market class. For example, Cantalupensis contains market classes such as Charentais, Shipper, Ogen, and Galia (Table 1).

Several simply inherited traits in melon dramatically effect plant phenotype (Pitrat, 1994). Fruit morphology is affected by the *O* (oval shape), *ri* (ridged surface), and *s* (sutures) genes. Fruit skin/flesh color is conditioned by color genes such as *gf* (green flesh color, recessive to salmon), *w* (white mature color), *Wi* (white immature color), and *Y* (yellow epicarp). Even though individual phenotypes within these market classes can be readily identified by visual inspection, the relative genetic distances (GDs) among market classes and between individual accessions within these groupings have not been rigorously defined. Such genetic identity information would be useful for germplasm management and plant variety protection in melon.

Molecular markers have been shown to be useful for diversity assessment in a number of plant species (Waugh & Powell, 1992; Bretting & Widrechner, 1995). Esquinas-Alcazar (1981) combined melons from 11 countries of origin into four groups using isozyme loci. Staub et al. (1997a) used 19 isozyme and 47 random amplified polymorphic DNA (RAPD) loci to characterize genetic differences among accessions in five melon groups (Cantalupensis, Conomon, Flexuosus, Inodorus, and Momordica). Phenotypic and molecular variation have also been used to characterize broad genetic intraspecific relationships (Stepansky et al., 1999) and narrow relationships

between selected landraces and melon lines (Mo-Suk et al., 1998).

Initial reports on DNA sequence variation in melons, however, have documented a relatively low number of base substitutions among melon cultivars (Shattuck-Eidens et al., 1990). Neuhuasen (1992) used restriction fragment length polymorphisms (RFLPs) for melon cultivar discrimination, and found that only 33% of the tested probes were useful in differentiating at least one of the seven melon accessions tested. The majority of the informative RFLPs detected only two hybridization patterns in a sample of 44 diverse melon accessions. In contrast, higher levels of polymorphism (71%) associated with simple sequence repeat (SSR) loci have recently been demonstrated in *C. melo* (Katzir et al., 1996). Gene diversity values obtained with SSRs in melon were high (0.42–0.75) with two to six alleles for each SSR in a sample of eight varieties belonging to four melon groups.

For routine use in germplasm assessment and management, a marker methodology must be technically simple, inexpensive, and amenable to high throughput sample processing. Where RAPDs and SSRs have been developed and characterized, they are particularly well suited to high through-put systems required for germplasm assessment because of their simplicity, speed, and relatively low cost (Williams et al., 1990; Staub et al., 1996a; Rafalski & Tingey, 1993; Wu & Tanksley, 1993; Beckmann & Soller, 1990; Davies, 1993; Diwan & Cregan, 1997). Garcia et al. (1998) used 115 RAPD loci and 24 agronomic traits to estimate genetic distances among 32 elite breeding lines to evaluate their potential as tools for germplasm management. Garcia et al. (1998) used a narrow germplasm base for their diversity assessment (one commercial source) focusing mainly on Galia and Piel de Sapo market classes. Since the GDs among several market classes of melons have not yet been rigorously defined and the value of SSR loci has not been evaluated in elite germplasm, we used RAPD and SSR marker systems in melon to: 1) evaluate the relative concordance among three GD estimators for discriminating among accessions in an array of market classes in the horticultural groups, Cantalupensis and Inodorus; 2) characterize genetic relationships among and between accessions in these groups, and; 3) determine the relative efficacy of these marker systems for diversity analysis in the populations examined (objectives). This diversity analysis provides GD benchmarks in several previously undocumented elite melon

Table 1. Plant habit, flowering and fruit characteristics of melon (*Cucumis melo* L. subspecies *melo*) groups separated into horticultural groupings.

Subspecies group <sup>e</sup>	Market class	Fruit diameter (cm)	Fruit weight (kg)	Fruit shape	Fruit flesh color	Fruit netting or stripping <sup>c</sup>	Relative days to fruit maturity	Relative canopy size <sup>b</sup>	Sex expression
Cantalupensis	Galia	10–12	0.8–1.2	Globular	Green	P	80–90	Large canopy	Strongly andromonoecious <sup>d</sup>
Cantalupensis	Ogen	10–12	0.7–1.2	Globular	Green	P	85–95	Medium canopy	Strongly andromonoecious
Cantalupensis	Charentais	10–12	0.8–1.2	Globular	Orange	A or P	85–95	Large canopy	Andromonoecious/monoecious
Cantalupensis	Eastern shipper	12–15	1.2–2.0	Globular	Orange	A/P	80–90	Large canopy	Andromonoecious
Cantalupensis	European shipper <sup>d</sup>	10–12	0.8–1.2	Globular	Orange	A or P	85–95	Large canopy	Strongly Andromonoecious
Inodorus	Honeydew	14–16	1.8–2.5	Globular	White	A	85–95	Large canopy	Andromonoecious
Inodorus	Cassaba (Rochet)	13–16	1.5–2.5	Elongate	White	A or P	85–95	Large canopy	Strongly andromonoecious
Inodorus	Cassaba (Piel de Sapo)	14–17	1.5–2.5	Ovoid	White	A	80–90	Large canopy	Andromonoecious
Inodorus	Cassaba (Yellow Canary)	14–17	1.5–2.5	Ovoid	White	A	80–90	Medium canopy	Andromonoecious
Conomon	—	5–6	0.8–1.5	Elongate	White	A or P	80–90	Medium canopy	Strongly andromonoecious
Flexuosus	—	2–4	0.5–0.8	Elongate	White	A	80–90	Large canopy	Andromonoecious/monoecious

<sup>a</sup> Andromonoecious = male and hermaphroditic flowers, and strongly andromonoecious = a greater proportion of hermaphroditic flowers.

<sup>b</sup> Relative size, where medium denotes canopies <1.0 m, and large = ≥ 1.0 m in diameter.

<sup>c</sup> A = absent, P = present, A/P = mostly absent, but some stripping can be detected, A or P = variable presence or absence.

<sup>d</sup> Although unique in shelf-life characteristics, fruit morphology can be either like Galia or Charentais horticultural types depending on market.

<sup>e</sup> According to Munger and Robinson (1991) and Whitaker & Davis (1962).

market classes, and presents useful information for germplasm management of melon.

## Materials and methods

### *Plant material*

Forty-six melon accessions were obtained from five seed companies and the United States Department of Agriculture, Agricultural Research Service (Table 2). Based on the information received from the donor, these accessions were classified as to their Subspecies *Melo* grouping (Cantalupensis, Inodorus, Flexuosus, or Conomon) according to Whitaker & Davis, 1962 and Munger & Robinson (1991), and then further partitioned into market classes for Cantalupensis (Char-entais, Shipper, European, U.S. western and U.S. eastern types, Galia, and Ogen) and Inodorus (Honey dew type, and Cassaba (syn. Spanish) Rochet, Piel de Sapo and Yellow Canari types) according to Bailey & Bailey (1976) and Munger & Robinson (1991) for genetic analyses. Some accessions (Nos. 19, 22, 38, 39, 43, 45, and 46) were designated as reference accessions and included as a subset of accessions in each analysis. These accessions were chosen because of their *Melo* subspecies classification [e.g., no. 39 (Flexuosus), no. 43 (Conomon)], U.S. market importance [no. 46 (Top Mark; U.S. western shipper)] and/or horticultural traits [nos. 19 and 22 (European Honey Dew), no. 38 (Canari), and no. 44 (U.S. western shipper)] according to Staub et al. (1997a). The Group Flexuosus and Conomon accessions chosen are representative of the diversity in these groups.

The DNA from 15 plants of each accession was bulked for analysis. Regardless of marker type, DNA was extracted from young leaf tissue of plants using a CTAB extraction procedure (Staub et al., 1996b; Maniatis et al., 1982). The DNA was then quantified on a Hoefer TKO 100 mini-fluorometer (Hoefer Scientific Instruments, San Francisco, California) following the manufacturer's protocol.

### *RAPD analysis*

The RAPD analysis was similar to that of Williams et al. (1990) with modifications for melon (Staub et al., 1996b). All polymerase chain reaction (PCR) solutions were purchased from Promega (Madison, Wisconsin) and PCR was accomplished according to Staub et al. (1997a). Each PCR had a volume of 15  $\mu$ l and contained 3.0 mM MgCl<sub>2</sub>, 0.2 mM dNTPs (0.05

mM of each: dATP, dGTP, dTTP and dCTP), 15 ng DNA, 0.3  $\mu$ M primer, commercial polymerase buffer and one unit Taq DNA polymerase. The 10-mer primers A1 to AX20 were purchased from Operon Technologies (Alameda, California), and the primers BC200 to BC699 were obtained from the University of British Columbia (Vancouver, BC, Canada). PCRs were conducted with the Perkin Elmer GeneAmp PCR System 9600 (Norwalk, Connecticut) thermocycler using the following cycling profile: 94 °C/ 4 min; 3 cycles of 94 °C/ 15 sec, 35 °C/ 15 sec, 59 sec ramp to 72 °C/ 75 sec; 40 cycles of 94 °C/ 15 sec, 40 °C/ 15 sec, 59 sec ramp to 72 °C/ 75 sec; 72 °C/ 7 min, and indefinite soak at 4 °C.

After completion of the PCR, 3  $\mu$ l of loading dye (0.25% bromophenol blue, 0.25% xylene cyanoll FF, 15% Ficol) was added to each reaction tube. The samples were electrophoresed in 1.6% agarose gels (20 x 25 cm) containing 0.5  $\mu$ g/ml ethidium bromide in TAE buffer [4.84% tris (trishydroxymethylamino-methane), 1.14% acetic acid, 0.375% EDTA] for three hrs at approximately 100 volts. The gels were immediately photographed using the Eagle Eye still video system (Stratagene, LaJolla, California).

Only consistent and heritable RAPD primers were used (unpublished data). A total of 135 RAPD marker loci (64 primers) were used for germplasm evaluation. Primers used were B12, C1, D7, F1, F4, G8, I4, I16, N6, L18, W7, AB14, AD12, AD14, AE6, AF7, AF14, AG15, AG18, AJ18, AK16, AL5, AM2, AN5, AO8, AO19, AT1, AT2, AT5, AT7, AT15, AU1, AU2, AS14, AV11, AW10, AW14, AX16, BC226, BC231, BC252, BC299, BC318, BC280, BC302, BC340, BC388, BC403, BC407, BC526, BC551, BC605, BC617, BC618, BC627, BC628, BC635, BC641, BC642, BC646, BC650, BC652, BC663, BC654. A marker was considered repeatable if PCR yielded a consistent result in all of three (or more) replications (Staub et al., 1997a).

### *SSR analysis*

The 17 (54 bands) SSR markers used in this study are presented in Table 3. Fourteen SSRs were constructed from a genomic DNA library that was prepared from cultivar Noy Yizre'el, and three SSRs were derived from EBML database sequences (Katzir et al., 1996; Danin-Poleg et al., in preparation).

The PCR amplification of SSR loci followed procedures described by Katzir et al. (1996) in which reaction mixtures for the PCR amplification of mi-

Table 2. Melon (*Cucumis melo* L.) germplasm used for diversity analysis.

Analysis code number	Seed source <sup>c</sup>	Source number	Subspecies <i>Melo</i> grouping	Horticultural grouping	Genetic type/name <sup>b</sup>	Analysis grouping <sup>a</sup>
1	RZ	1	Cantalupensis	Galia	IL	1
2	RZ	2	Cantalupensis	Ogen	IL	2
3	RZ	3	Cantalupensis	Galia	IL	1
4	RZ	4	Cantalupensis	Ogen	IL	2
5	RZ	5	Cantalupensis	Ogen	IL	2
6	RZ	6	Cantalupensis	Ogen	IL	2
7	RZ	7	Cantalupensis	Charentais	IL	3
8	RZ	8	Cantalupensis	Charentais	IL	3
9	RZ	9	Cantalupensis	Charentais	IL	3
10	RZ	10	Cantalupensis	Galia	IL	1
11	LM	23840	Cantalupensis	Charentais	OP, 'Printadoux'	3
12	LM	23841	Cantalupensis	Ogen	IL	2
13	LM	23842	Cantalupensis	Ogen	IL	2
14	LM	23846	Cantalupensis	Galia	IL	1
15	LM	23847	Cantalupensis	European shipper	IL	4
16	LM	23848	Cantalupensis	European shipper	IL	4
17	LM	23849	Cantalupensis	European shipper	IL	4
18	LM	23850	Cantalupensis	European shipper	IL	4
19	LM	23851	Inodorus	Honey dew	IL	RA
20	LM	23852	Cantalupensis	European shipper	IL	4
21	LM	23853	Cantalupensis	European shipper	IL	4
22	LM	24138	Inodorus	Honey dew	OP, 'Witte Suiker'	RA
23	Zu	M-700	Cantalupensis	Charentais (netted)	IL	3
24	Zu	M-473	Cantalupensis	Galia	IL	1
25	Zu	M-442	Inodorus	Cassaba (Rochet)	IL	5
26	Zu	M-103	Inodorus	Cassaba (Rochet)	IL	5
27	Zu	M-126	Inodorus	Cassaba (Piel de Sapo)	IL	5
28	Zu	MP-136	Inodorus	Cassaba (Yellow Canari)	IL	5
29	Zu	MY-124	Inodorus	Cassaba (Yellow Canari)	IL	5
30	Zu	OP-Canario	Inodorus	Cassaba (Piel de Sapo)	OP	5
31	Zu	OP-Requete	Cantalupensis	Charentais	OP	3
32	Zu	M-203	Cantalupensis	Galia	OP, 'Mango'	1
33	Peto	PSX 17291	Inodorus	Cassaba (Yellow Canari)	IL	1
34	Peto	Bardino	Inodorus	Cassaba (Yellow Canari)	F1	5
35	Peto	Etoile	Cantalupensis	Charentais	F1	3
36	Peto	PSX 105292	Cantalupensis	U.S. eastern market	IL	4
37	Peto	Yuma	Cantalupensis	Galia	F1	1
38	Peto	Dorado	Inodorus	Cassaba (Yellow Canari)	F1	RA
39	Peto	Snake melon	Flexuosus		NA	RA
40	Peto	Elton	Cantalupensis	U.S. eastern market	F1	4
41	Peto	Pulsar	Cantalupensis	U.S. eastern market	F1	4
42	Peto	Durango	Cantalupensis	U.S. western shipper	F1	4
43	Peto	Freeman cucumber	Conomon		NA	RA
44	HM	HMX-95-84	Cantalupensis	U.S. western shipper	IL	4
45	HM	HMX 2608	Cantalupensis	U.S. eastern market	IL	RA
46	USDA	Top Mark	Cantalupensis	U.S. western shipper	OP	RA

<sup>a</sup> RA = reference accession used in all analyses.<sup>b</sup> IL = inbred line, OP = open pollinated variety, F1 = single cross hybrid, NA = not applicable.<sup>c</sup> RZ = Rijk Zwaan Seeds De Lier, The Netherlands, LM = Leen de Mos BV, Granvendzade, The Netherlands, Zu = Zaadunie BV, Enkuizen (now Novartis), Peto = Peto Seed Company (now Seminis), Woodland, CA, HM = Harris Moran Seed, Modesto, CA, and USDA = United States Department of Agriculture, Agricultural Research Service, Salinas, CA.

Table 3. Description of single sequence repeat marker alleles (17 loci) detected among 46 melon (*Cucumis melo* L.) accessions.

SSR designation <sup>a</sup>	Core motif	Number of alleles among 46 accessions	Expected size (bp) <sup>b</sup>	Chromosomal position <sup>c</sup>
From genomic library				
Melon				
CMTC13	(TC) <sub>12</sub> (CG) <sub>5</sub> (AG) <sub>3</sub>	4	92	
CMAG59	(GA) <sub>2</sub> A(AG) <sub>8</sub>	3	124	E
CMGA127	(GA) <sub>13</sub> A(GA) <sub>2</sub>	3	138	
CMGA128	(GA) <sub>10</sub> AA(GA) <sub>2</sub>	4	119	J
CMGA15	(GA) <sub>7</sub>	2	150	F
CMCT44	(CT) <sub>10</sub> TGTT(CT) <sub>3</sub>	2	104	
CMTC47	(TC) <sub>9</sub> (CT) <sub>6</sub>	3	168	5
CMGA104	(GA) <sub>14</sub> AA(GA) <sub>3</sub>	6	125	
CMGT108	(GT) <sub>9</sub> N65(CT) <sub>7</sub>	3	187	
CMAT141	(AT) <sub>7</sub> GT) <sub>6</sub>	2	176	E
CMCTT144	((CTT) <sub>10</sub> CTAC(CTT) <sub>4</sub>	7	192	E
CMCCA145	(CCA) <sub>5</sub>	2	142	G
CMACC146	(ACC) <sub>9</sub>	2	152	E
CMTC160a+b	(TC) <sub>2</sub> (TCC) <sub>2</sub> (CT) <sub>8</sub> N122(TC) <sub>8</sub>	2	215	6
From database				
Melon				
CMAT35	(TA) <sub>3</sub> AA(TA) <sub>2</sub> C(AT) <sub>7</sub>	3	110	
Cucumber				
CSGA057	(GA) <sub>8</sub>	3	211	
CSCCT571	(CCT) <sub>5</sub> CTT(CT) <sub>2</sub>	3	209	C

<sup>a</sup> SSR markers described by: Katzir et al. (1996). and Danin-Poleg et al. in preparation.

<sup>b</sup> Expected size of the amplification product for each SSR locus estimated from library directly or according to the EBML database sequence.

<sup>c</sup> Danin-Poleg, Reis, Baudracco-Arnas, Pitrat and Katzir, submitted.

cro-satellite loci contained 60 ng of plant genomic DNA, 1 mM of Mg<sup>2+</sup>, 8–10 pmole of 3' and 5' primers, 166 µM of dATP, dTTP, dGTP, 2 µM of dCTP, 0.1 µl of 3000 Ci/mmol [ $\alpha$ -<sup>32</sup>P] dCTP, 1 x Taq Buffer (Advanced Biotechnologies, UK), 1 unit of Taq DNA polymerase (Advanced Biotechnologies, UK), in a total volume of 15 µl. The amplification program was as follows: 30 sec denaturation at 94 °C, 30 sec annealing at 51 °C, and 60 sec extension at 72 °C for 34 cycles on a thermocycler (PTC-100 MJ Research Inc.). PCR products (3.0 µl/lane) were separated on a DNA sequencing gel, containing 6% polyacrylamide, 8M urea and 1 × TBE, at 60 W constant power for 1.5–2.5 h. The sequencing reactions of the four nucleotides of M13 ssDNA were used as molecular-weight standards to determine the exact nucleotide length of the denatured PCR products. After drying, the gels were exposed to a Kodak XAR-5 film (Eastman Kodak).

### Data analysis

To describe genetic relationships, RAPD and SSR marker data were used to calculate GD estimates between and among accessions using three estimators: simple matching (Sokal & Sneath, 1963), Nei's distance-D (Nei, 1973, 1987), and Jaccard's coefficient (Debener, 1990). A distance estimate was calculated for each pair-wise comparison among the 46 accessions for each marker type. RAPD marker band phenotypes were designated as '+' (present) or '-' (absent). Because of the codominant nature of SSR markers, intra-locus variation can be scored as + ('++') or - ('--') or +/- ('-+', heterozygote), and such information is useful in differentiating homozygous and heterozygous individuals. The allelic frequency of individual accessions, however, could not be calculated because the DNA sample from each accession was a bulked sample (i.e., a pooled marker profile). Thus, repeatable RAPD banding phenotypes were scored

as presence (+) or absence (–) of a DNA band, and assigned 1 or 0 for analysis, respectively. For SSR variation, alternate homozygotes were assigned as 1 or 0, and heterozygotes were given the value of 0.5 for analysis.

The number and relative mobility of alleles affects the calculation of GD. For most accessions a maximum of two alleles were detected by SSR analysis. Many of these accessions were F1 hybrids, and thus the presence of two alleles was predicted based on the homozygous and homogenous nature of their inbred components. Only the DNA of Flexuosus possessed three alleles when investigated with two melon SSR markers (CSGA057 and CMTC13). The predominant presence of two alleles at putative loci in both marker types allowed for comparative analysis.

Simple matching coefficients ( $S_{ij}$ ) were calculated for each comparison as  $s_{ij} = (a+d)/n$ , where  $a$  = the number of bands in common to both accessions ('++' matches),  $d$  = the number of missing bands in common to both accessions ('–' matches), and  $n$  = the total number of bands (includes '++', '–', '+–', and '+–') (Sokal and Sneath, 1963). The simple matching coefficient by definition may vary from 1 to 0, where 1 indicates that a pair has identical banding patterns and 0 indicates that a pair has completely opposite banding patterns.

Jaccard's coefficients were calculated  $J_{ij} = a/(a+c)$ , where  $a$  = the number of bands in common to both accessions ('++' matches) and  $c$  = the number of missing bands in one accession, but present in the other ('–' and '+–' matches) (Debener, 1990). Like the simple matching coefficient, the Jaccard's coefficient can vary from 1 to 0, where 1 indicates that a pair has identical banding patterns and 0 indicates that a pair has uniformly contrasting (opposite) banding patterns. The major difference between these two similarity measures is that, unlike the Jaccard's coefficient, '–' matches are used in the calculation of simple matching coefficient values. The similarity measurements ( $S_{ij}$  and  $J_{ij}$ ) were converted into a GD measurement by calculating the complement of each coefficient ( $1-S_{ij}$  and  $1-J_{ij}$ ) as described by Spooner et al. (1996).

Nei genetic distances were calculated as  $d_{ij} = \ln[(\sum_k X_{ki} X_{kj}) / \sqrt{(\sum_k X_{ki}^2 X_{kj}^2)}]$ , where  $X_{ki}$  and  $X_{kj}$  are the frequencies of the  $i$ th and  $j$ th alleles at the  $k$ th locus, respectively. Thus, the values of  $X_{ki}$  and  $X_{kj}$  can be either 1 or 0. As with Jaccard's coefficient, Nei GD estimation puts no weight on the '–' matches. The application of Nei's distance calculation to dominant marker systems results in the classifica-

tion of all accessions [either a presence (AA, Aa) or absence (aa)] as homozygous (AA or aa). The square of '++' matches is 1, and the square of '–' is zero (no weight). The numerator thus becomes the sum of the '++' ( $1 \times 1$ ) matches (i.e., the '+–' =  $1 \times 0$  and drop out of the calculation), and the denominator reflects the square root of the sum of the squares of the cross products of the '++' (1) or '+–' (0) matches squared.

Cluster analysis and multidimensional scaling (MDS) of GD estimates were used to analyze genetic relationships. The unweighted pair-group method using the arithmetic average clustering procedure (UPGMA) was employed, and the resulting dendrograms were constructed with the genetic distance matrix using the computer program NTSYS-pc version 1.80 (Rolf, 1997). MDS was employed to provide a pictorial representation of relationships among accessions within a market class and reference accessions (Table 2). These hierarchical procedures allowed for biologically consistent simplification of the data. The MDS procedure assigns a point in three-dimensional space to each accession based on the relationship of each pairwise GD value between any two accessions (Wilkinson, 1989). MDS attempts to minimize the 'stress' between points during multi-dimensional scalar analysis, and provides a goodness of fit statistic between 0 and 1 in which values near 0 indicate better fit (i.e., minimal tension between points in hyperspace).

The degree of concordance among the GD estimators was determined by visual appraisal of graphic depictions of GD after MDS, and correlation analysis (objective 1). A nonmetric MDS scaling of each of the six similarity matrices (2 marker types  $\times$  3 GD estimators) was performed initially to investigate the similarities between GD estimators. Spearman rank correlation coefficients ( $r_s$ ) were calculated for all pairwise comparisons between marker types and GD estimators (2 marker  $\times$  3 estimators) using all accessions in the initial analysis (Steel & Torrie, 1980).

The estimator giving the most dispersion and having the highest concordance with other estimators was identified, and then MDS was applied to characterize relationships among and between accessions using the selected GD estimator (objective 2). The statistical significance for the independence of five identified market classes (Table 2) was tested according to Staub et al. (1997b). To compare the two groups of accessions for the frequency of RAPD and SSR markers (number of accessions possessing a marker / total

number of accessions in the group), a two-sample chi-square goodness of fit test was employed (Gibbons, 1976). The expected value of marker frequency within each group was calculated based on the frequency of accessions possessing a RAPD or SSR marker among all accessions.

To make overall comparisons between any two groups (i.e., marker or market class type test groups), the results of individual chi-square tests were pooled using the procedure of Sokal & Rohlf (1981) for combining probabilities of independence tests. Probability ( $P$ ) values were calculated separately for each chi-square test for the 135 RAPD and 54 SSR bands. An overall probability that the two groups of accessions being compared had identical marker frequencies was calculated by summing the natural log of all  $P$  values. Since the distribution of this sum closely follows the chi-square distribution, an overall probability can be calculated.

#### *Marker efficiency comparisons*

In order to determine the efficiency of each marker type per unit information (band) (objective 3), the variance within each marker data set (SSR and RAPD) was empirically estimated using a bootstrap sampling procedure (Efron & Tibshirani, 1986). A subset of a given number of  $N$  polymorphic bands were generated. The  $N$  bands were selected at random from the entire set of possible bands for each data set. Sampling was done with replacement, thus allowing for the probability of  $1/N$  of bands per data set to be selected at any one time from the data set. The band usage frequency was continuously monitored to detect any bias in the use of bands resulting from a total of 1000 subsamples (bootstraps). The coefficient of variation (CV) was used to determine the magnitude of the variance because there is a linear relationship between genetic variances and means of distance estimates (Tivang et al., 1994). Comparisons among marker types can be made since differences in the variances of each data set can be normalized.

## **Results and discussion**

#### *Concordance among genetic distance estimators*

Genetic relationships based upon variation in RAPD and SSR profiles among and between the accessions examined are depicted in Figures 1 and 2. Stress values associated with the scalar depiction of the

data sets after MDS were low (0.14 to 0.17) regardless of distance estimation method, indicating good fit of the data in three dimensional hyperspace (Figure 1). Variation observed with amplification using RAPD primers AF7, AG15, AJ18, AL5, AN5, AS14, AU2, AV11, AW10, AX16, B12, D7, G8, N6, W7, BC231, BC318, BC388, BC469 BC652, and BC654 was important ( $p=0.01$ ) in the detection of genetic differences among and between accession groups (35% of variation explained). Likewise, SSR primers CMT13, CMT47, CMAG59, CMGA104, CMGT108, CMCTT144, and CSGA057 were important ( $p=0.01$ ) discriminators for the detection of inter- and intra-group differences (24% of variation explained).

Different methodologies have been proposed for genetic diversity analysis using dominant and codominant genetic markers (Noli et al., 1997; Akagi et al., 1997; Xiao et al., 1996). The use of simple matching for the calculation of genetic distance is not recommended for SSR marker data because the absence of a specific band indicates the absence of the corresponding allele whereas alternate alleles may be detectable in other genotypes in the same population at any given locus. The number and abundance of alleles present depends on the species and the variation present in the population being examined.

It is important to identify differences in GD estimates among estimation procedures. Such differences, if significantly large, may dramatically affect the interpretation of results. Concordance of the MDS plots was assessed by the degree of correlation between the GD estimators (Jaccard's coefficient, simple matching coefficient, and Nei's distance-D) and the markers (RAPD and SSR). Genetic distance estimates derived by the three estimation methods were highly ( $p > 0.0001$ ) correlated ( $r_s$  0.64 to 0.99) when comparisons were made between methods within a particular marker system (e.g., simple matching RAPD by Jaccard's coefficient RAPD). For instance, correlations between combinations of simple matching coefficient, Nei's D, and Jaccard's coefficient using RAPD data were high ( $r_s > 0.98$ ). However, correlations between combinations of these estimators using SSR data were comparatively lower ( $r_s$  of simple matching  $\times$  Nei's D = 0.64; simple matching  $\times$  Jaccard's coefficients = 0.77; Jaccard's coefficient  $\times$  Nei's D = 0.94). Lower correlations among SSR data sets is expected due to the fact that simple matching is not recommended for SSR marker data. Another factor that might have affected correlations between GD estimators is the dif-



Figure 1. Genetic distance (Jaccard's coefficient) relationships among different market classes of *Cucumis melo* L. accessions as depicted by multi-dimensional scaling of variation observed with 64 RAPD primers (reference accessions in bold; Con = Conomon, Flex = Flexuosus, CC = Cassaba-Yellow Canari, CH = Cassaba-Honeydew, USW = U.S. Shipper, USE = U.S. Eastern Market).

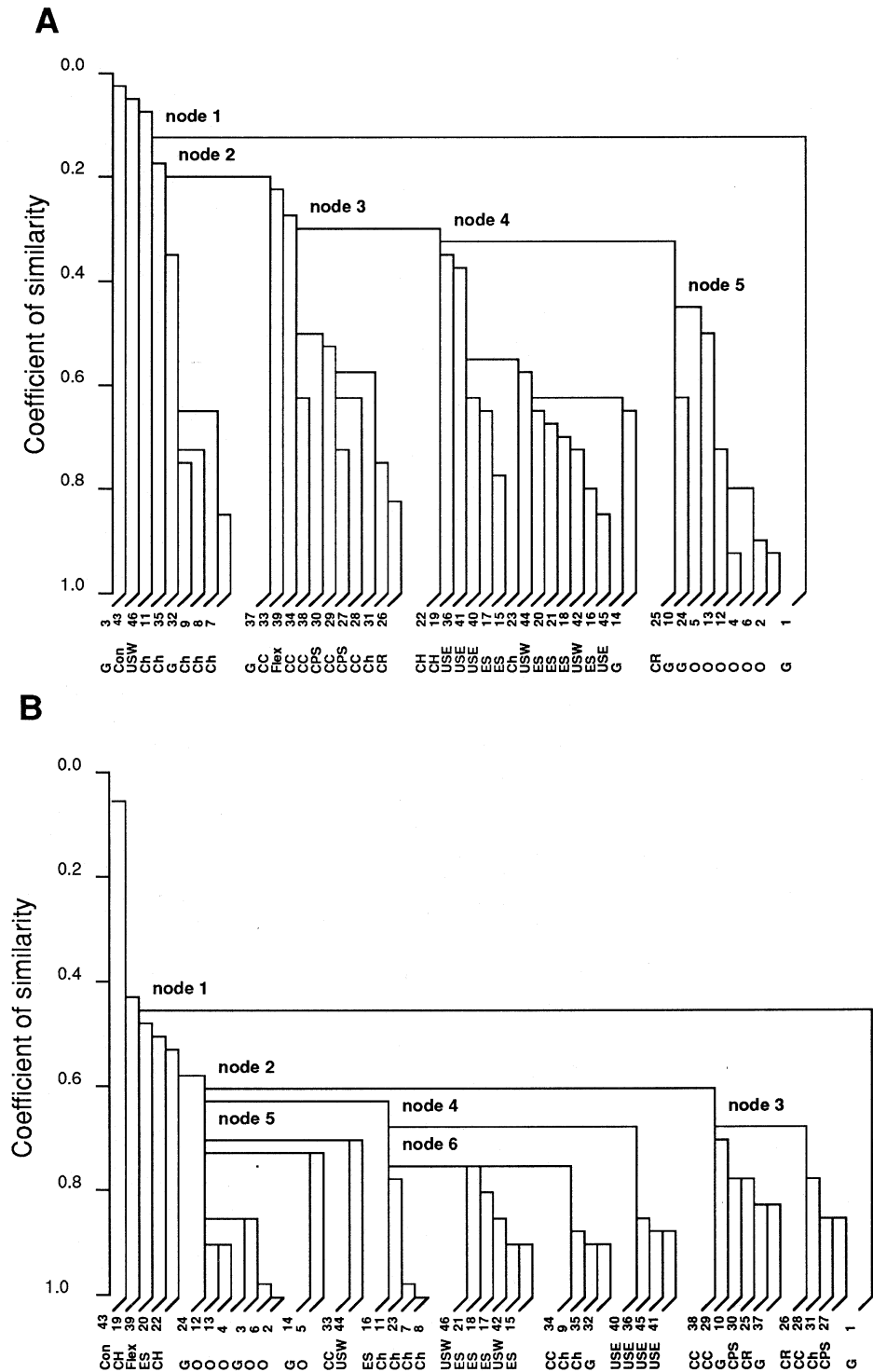


Figure 2. Cluster analysis (by UPGMA) of 46 *Cucumis* accessions grouped using genetic distances (Jaccard's coefficient) as estimated by (A) 154 RAPD bands and (B) 54 SSR bands as framing criteria (Con = Conomon, Flex = Flexuosus, g = Galia, Ch = Charentais, CC = Cassaba-Yellow Canari, CPS = Cassaba-Piel de Sapo, CR = Cassaba-Rochet, CH = Cassaba-Honeydew, ES = European Shipper, O = Ogen, USW = U.S. Shipper, USE = U.S. Eastern Market).

ference in the number of putative loci compared by each marker [SSR markers (17 loci) vs. RAPDs (64)].

In contrast, correlations between a particular marker system and a GD estimator (e.g., Jaccard's coefficient RAPD and Jaccard's coefficient SSR) were lower ( $r_s = 0.17$  to  $0.40$  depending on coefficient used;  $p > 0.001$ ). The low correlations observed between marker systems when comparing GDs derived from a particular estimator (Table 4; Figure 2) is likely due to the disparity in number of markers used within a system (54 SSR bands vs. 135 RAPD bands). While GD estimations using SSR markers ranged between 0.0 (all markers identical; no. 2  $\times$  no. 6 and no. 7  $\times$  no. 8) to 0.97 (no. 2  $\times$  no. 43, no. 6  $\times$  no. 43, no. 20  $\times$  no. 43), estimations employing RAPD markers ranged between 0.04 (no. 12  $\times$  no. 13) to 0.65 (no. 43  $\times$  no. 3).

Simple matching and Jaccard's coefficients are rather simplistic and similar in their calculation (i.e., both have similar metric properties) (Jackson et al., 1989). Nevertheless, the numerator of these coefficients is different ( $a + d$  vs.  $a + c$ ) such that the behavior of these similarity coefficients may be data specific (Janowitz, 1980; Hubalek, 1982). This was the case in the present analysis where mean GD values calculated from variation at RAPD loci were  $0.27 \pm 0.08$  and  $0.31 \pm 0.09$  for simple matching and Jaccard's coefficients, respectively.

Concordance between Jaccard's coefficient and Nei's D estimates of GD might be predicted to be high since both estimators give no weight to '—' (value of 0) matches. This was the case in this study where correlation between GD estimation using Jaccard's coefficient and Nei's D was 0.99 ( $p > 0.0001$ ). Mean Jaccard's coefficient GD values were higher ( $0.31 \pm 0.09$ ) than those obtained by estimation using Nei's D value calculation ( $0.21 \pm 0.08$ ) suggesting the conservative nature of the Nei's D estimator (Jackson et al., 1989) even when applied to the RAPD data set (i.e., dominant marker providing no allelic information) of this study.

Staub et al. (1997a) used 43 RAPD loci to differentiate Group Conomon and Flexuosus group accessions from Group Cantalupensis and Inodorus accessions. They found that cluster analysis grouped Conomon and Flexuosus accessions into a separate clade from Cantalupensis and Inodorus accessions. The subsp. *agrestis* Group Conomon and Flexuosus accessions employed in our study were the same as those used by Staub et al. (1997a). We found that these accessions examined were genetically distinct ( $p > 0.001$ ),

and different ( $p > 0.001$ ) from the Cantalupensis and Inodorus accessions analyzed, thus recapitulating results of Staub et al. (1997a). The minimum and maximum GD between the Flexuosus accession (no. 39) examined and any other accession was 0.11 (no. 39  $\times$  no. 34; simple matching SSR) and 0.57 (no. 39  $\times$  no. 11; Nei's D RAPD), respectively. The minimum and maximum GD between the Conomon accession (no. 43) examined and any other accession was 0.31 (no. 43  $\times$  no. 23; Nei's D RAPD) and 0.99 (no. 43  $\times$  no. 20; Nei's D SSR), respectively.

#### *Genetic relationships among and between accessions in market classes*

Based on the concordance of the GD estimators examined, the simplicity and minimal assumptions leading to GD estimation using Jaccard's coefficient (Jackson et al., 1989), and the use of this estimator by Garcia et al. (1998), Jaccard's coefficient was used for all subsequent comparisons. This allowed for the interpretation of genetic relationships (Figures 1 and 2; Table 4), and the comparison of these relationships with those of Garcia et al. (1998).

The difference in numbers of bands used to detect genetic variation does not allow for specific comparisons between marker data sets in this study. Nevertheless, the genetic relationships identified using these markers were generally similar. In both marker cluster analyses, Galia accession no. 1 differed markedly from the other accessions examined (Figure 2, panels A and B). Cluster analysis using RAPD markers resulted in the partitioning of the subsp. *agrestis* Group Conomon (no. 43; node 1) accession to reveal its unique relationship to the *C. melo* accessions examined (Figure 2, panel A). The Flexuosus group accession (no. 39; node 3) clustered with the Inodorus accessions surveyed indicating shared RAPD banding profile similarities. Ogen (node 5), Shippers (Western and European; node 4), and Cassaba (node 3) were grouped into distinct dendrogram positions. The Honeydew accessions (nos. 19 and 22; originating from same source) examined are considered Cassaba market types. However, these accessions grouped with Shipper types suggesting that their genetic affinity is closer to Shipper than to Cassaba types. While Galia accessions were dispersed throughout the tree (nodes 1–5), all Charentais accessions, except nos. 23 (node 4) and 31 (node 3), clustered in node 2. The variable nature of the Galia accessions examined is consistent with its unique pedigree (see below).

**Table 4.** Pairwise genetic distance values (Jaccard's coefficient) between melon genotypes included in several market classes using RAPD (below diagonal) and SSR (above diagonal) markers

[illegible]

Diversity analysis using SSR marker data produced a dendrogram with six main branching nodes (Figure 2, panel B). Accessions were partitioned into specific nodes based on market class. With the exception of accession no. 34 (Peto Seed Company; node 6), all Cassaba types grouped together (node 3). Cantalu-pensis accessions grouped together and showed partial consanguinity with European shipper accessions (except no. 20), a Cassaba (no. 34), a U.S. eastern market type (no. 45), and with most U.S. western shipper types. Although the European and U.S. western shipper market types examined were similar, they were, in the main, partitioned into separate branches (nodes 4 and 6). Likewise, Ogen and Galia accessions were similar (node 5), and these accessions showed affinity with one U.S. western shipper accession (no. 44; Harris Moran Seeds). However, Galia accessions nos. 10 (Rijk Zwaan Seeds) and 37 (Peto Seed Company) were more similar to the Cassaba types examined (node 3) than other commercial Galia types. In parallel with RAPD analysis (panel A), SSR analysis identified the unique genetic identity of one Galia (no. 1) and the subsp. *agrestis* Group Conomon accessions examined. In contrast to RAPD analysis, analysis by SSR identified the Flexuous group accession and the Inodorus (Honey dew type) accessions (no. 22 and 19) as distinct from the other accessions examined.

Since significant positive correlations ( $p < 0.05$ ) exist between GD estimates using SSR and RAPD marker loci (above) and the analysis of RAPD variation resulted in GD estimates with lower variation (i.e., standard errors of mean GDs ranged between  $\pm 0.04$  to  $0.09$ ) than that using SSR variation ( $\pm 0.07$  to  $0.12$ ), specific GD comparisons are hereafter restricted to estimates from RAPD marker data sets. The mean GD among accessions of Ogen, Galia, Cassaba, Charentais, European shipper, and U.S. western shipper accessions employing RAPDs was  $0.11 \pm 0.04$ ,  $0.33 \pm 0.09$ ,  $0.21 \pm 0.04$ ,  $0.26 \pm 0.10$ ,  $0.17 \pm 0.05$  and  $0.22 \pm 0.08$ , respectively (Table 4; Figure 1). Mean GD estimates among Yellow Canari and Rochet accessions examined were  $0.27$  and  $0.19$ , respectively. When shipper types were taken collectively, mean GD among accessions was  $0.20 \pm 0.05$ . These results indicate that the genetic diversity among the Ogen accessions examined is limited when compared to Galia and Charentais accessions (Figures 1 and 2). The relatively narrow genetic diversity found in the Ogen accessions is consistent with the fact that this market class originated from open-pollinated individuals restricted to a narrow geographic range.

Likewise, the genetic diversity among Cassaba and Shipper (U.S. and European types taken collectively) types was similar.

Garcia et al. (1998) examined 14 Galia and 13 Piel de Sapo, one Rochet and one Yellow Canari accession. GD estimates given by Garcia et al. (1998) among Galia ( $0.41 \pm 0.08$ ) and Cassaba ( $0.33 \pm 0.09$ ; all types) market types were higher than those estimated in this study. While the GD between Rochet and Yellow Canari estimated by Garcia et al. (1998) was  $0.52$ , the GD between these types in our study was  $0.21$  (Table 4). These disparities could be due to the difference in the germplasm examined, and/or the number and type of primers used.

Among Ogen accessions, nos. 2 and 6 and nos. 12 and 13 were found to be most similar ( $GD = 0.05$ ), and nos. 12 and 5 were most dissimilar ( $GD = 0.17$ ) (Figure 1). Genetic distances among Galia accessions nos. 14, 24, 32 and 37 were relatively small ( $GD = 0.24$  to  $0.27$ ), while the GD between nos. 3 and 1 ( $GD = 0.56$ ) were relatively large. Among Cassaba accessions, nos. 29 and 30 and nos. 26 and 28 were most similar ( $GD = 0.14$ ), and no. 25 and 34 and nos. 25 and 29 were most distinct ( $GD = 0.27$ ). Charentais accession nos. 8 and 7 were most similar ( $GD = 0.09$ ) and nos. 8 and 31 and 7 and 31 were most dissimilar ( $GD = 0.37$ ). Within the European shipper accessions, nos. 15 and 18 ( $GD = 0.14$ ) and nos. 16 and 18 ( $GD = 0.10$ ) were the most similar, while nos. 15 and 20 were most distinct ( $GD = 0.28$ ). U.S. western shipping accessions nos. 42 and 44, and U.S. eastern market nos. 40 and 41 were most similar ( $GD = 0.12$ ), and nos. 36 and 46 and 41 and 46 were distant ( $GD = 0.37$ ).

Fruit size, shape and net type distinguish U.S. western and eastern types. While the epidermis of U.S. eastern types is ribbed and covered in a light, moderate netting, and U.S. western shipping types have a dense and thickly netted epidermis without ribs or vein tracts, these differences are controlled by relatively few genes (Pitrat, 1994). The relatively close GDs ( $GD$  ranged between  $0.15$  to  $0.21$ ) among the U.S. western and eastern types examined might have been predicted based on common ancestral germplasm.

Significant genetic differences between all market classes were detected ( $p > 0.001$ ). The GDs of pairwise comparisons between Galia and other melon market types were always the largest (Table 4). The mean GD between Galia and Cassaba, Charentais, European shipper and U.S. shipper (western and eastern taken collectively) types was  $0.34 \pm 0.09$ ,  $0.37 \pm 0.10$ ,  $0.34 \pm 0.10$ , and  $0.35 \pm 0.11$ , respectively. Significant dif-

ferences ( $p > 0.001$ ) in GD between Galia accessions no. 1 and 3 (pedigrees proprietary) and accessions in other market classes define inter-market class differences, and indicate that Galia germplasm can differ substantially from other market classes. The mean GD between Galia no. 1 and Ogen, Cassaba, Charentais, and European and U.S. shipper types was  $0.40 \pm 0.01$ ,  $0.43 \pm 0.03$ ,  $0.45 \pm 0.02$ ,  $0.46 \pm 0.02$ ,  $0.48 \pm 0.03$ , respectively. The mean GD between Galia no. 3 and Ogen, Cassaba, Charentais, and European and U.S. shipper types was  $0.40 \pm 0.02$ ,  $0.51 \pm 0.03$ ,  $0.52 \pm 0.02$ ,  $0.51 \pm 0.03$ ,  $0.51 \pm 0.04$ , respectively. Likewise, in each market class comparison, contrasts involving the U.S. western shipper cultivar Top Mark (no. 46) produced the greatest GDs (mean GD in group tests =  $0.42 \pm 0.06$ ) between the accessions examined. The mean GD between Ogen and Charentais types was also relatively great (GD =  $0.36 \pm 0.07$ ). These data suggest that broad genetic differences can exist between melon market classes.

Nevertheless, differences between market classes can be small, indicating limited genetic diversity between such classes (Table 4; Figure 2). For instance, comparisons between Ogen and European shipper (GD =  $0.29 \pm 0.03$ ), Ogen and U.S. shipper (GD =  $0.30 \pm 0.06$ ), Ogen and Galia ( $0.30 \pm 0.07$ ) types yielded the smallest mean GDs across market classes. Although such comparisons suggest that the Ogen accessions examined have similarities to each of these groups, specific comparisons indicate that such general inferences are inadequate given the accession differences observed. For instance, Ogen accessions were distant from the U.S. western shipper accession 'Top Mark' (no. 46) (mean GD =  $0.43 \pm 0.02$ ). Likewise, GDs involving pairwise comparison of some Galia accessions with accessions of other groups can also be relatively small. For instance, the GD between Galia no. 14 and European (mean GD =  $0.18 \pm 0.03$ ) and U.S. (mean GD =  $0.21 \pm 0.06$ ) shipper types was comparatively small.

Differences were also detected between European and U.S. shipping types ( $p > 0.001$ ) (Figure 2). However, the relative lack of genetic diversity among some shipper types was not predicted based on morphological differences and adaptation to different markets (Tables 1 and 4; Figure 1). For instance, the GD between nos. 16 (European) and 42 (U.S.) was relatively small (GD = 0.07) compared to that of nos. 17 (European) and 46 (U.S.) (GD = 0.31). Although these data suggest that some European and U.S. shipping types may have common parentage, extensive pedi-

gree information would be required to elucidate these apparent similarities.

#### *Relative efficacy of RAPD and SSR markers*

Since random sampling of a population results in a sample variance that is not equal to the value that would be obtained if the entire population were sampled, GD values are only approximations of absolute values. Greater precision in the estimation of GD values can be obtained by sampling more individuals in a target population and/or increasing the number of loci used in diversity assessment (Tivang et al., 1994). Analysis costs (e.g., number of samples needed to adequately estimate GD and the estimation methodology used) are an important consideration in germplasm management because of limited resources. Estimates regarding lower limits of the number of marker loci required for resolution of germplasm pools would be useful.

Examination of the relationship of coefficient of variation (CV) and the sample size (number of bands) can be useful in determining the variation associated with the estimation of genetic differences (Tivang et al., 1994). Estimation of genetic relationships become more definitive (i.e., lower CVs) as variances are reduced by sampling larger population sizes. Staub et al. (1997a) analyzed genetic diversity among genotypes drawn from a diverse array of subsp. *melo* (Cantalupensis and Inodorus) and *agrestis* (Conomon, Flexuosus, and Momordica) accessions using isozyme and RAPD loci, and then employed bootstrap sampling procedures to examine sampling variation associated with genetic difference estimations. Bootstrap analysis of marker loci indicated that the addition of more than 35 marker loci (isozyme or RAPD) did not decrease the mean CV leveled below about 33%. They concluded that 35 marker loci was a good minimum for diversity analysis among *C. melo* germplasm.

Empirical estimations of variances associated with each marker type in the accessions examined indicated that, per band, lower coefficients of variation can be attained in the estimation of GD when using RAPDs compared to SSRs (Figure 3). The disparity between the marker analyses made may be related to the amount of genome coverage characteristic of a particular marker system and its efficiency in sampling variation in a population (Staub et al., 1997a). Although the genomic position of some of the SSRs markers used is known (Table 3), the position of putat-

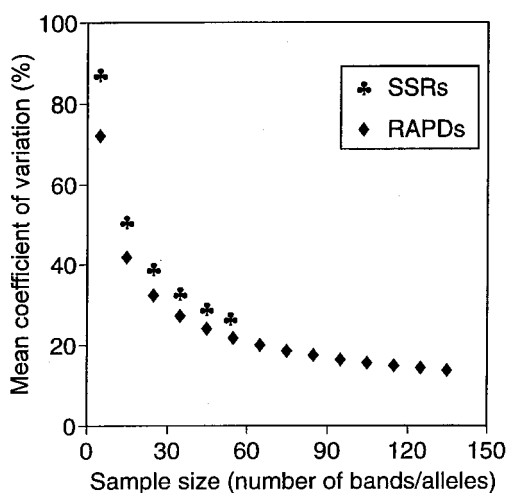


Figure 3. Sample variance of genetic distance estimation for *Cucumis melo* L. accessions as depicted as the relationship between the mean coefficient of variation (%) and the sample size [number of RAPD and SSR bands (alleles)] derived from a bootstrap procedure (analysis by J. Villand).

ive RAPD loci used in this has not been documented by isolation and hybridization experiments. However, RAPD loci in melon have been mapped (Baudracco-Arnas and Pitrat, 1996), and the loci used in our study show Mendelian segregation [3 (band present): 1 (band absent)] (Staub, unpublished data). It is not possible to determine the relative sampling efficiency per band (locus) relative to each marker type because information about total genomic coverage for both markers is inadequate. Nevertheless, even with the lower number of SSR bands examined, RAPD and SSR markers revealed similar genetic relationships among commercial market classes. A comparison of inter-marker heterozygosity (%) could not be made since all the RAPD markers used were dominant.

Bootstrap analysis indicated that the CVs from SSR and RAPD marker population resampling experiments begin to level at about 60 marker loci; the rate of decrease is comparatively minimal beyond 80 marker loci (Figure 3). These data suggest that for this germplasm array, 80 marker loci are adequate (~ CV of 17%; Figure 3) for assessing the genetic variation present in the population (Figures 1 and 2). These results and those of Staub et al. (1997a) indicate that 35 to 80 marker loci are an appropriate number for assessments of melon germplasm depending market type and relatedness.

Stepansky et al. (1999) used morphological characteristics of Naudin (1859) to partition 54 melon

accessions into several 'horticultural varieties.' Mo-Suk et al. (1998) used eight polymorphic RAPD bands to differentiate 52 Korean landraces and lines into two distinct groups. The melon lines examined could be further grouped into two subgroups (net and non-netted fruit types) using four RAPD markers. Likewise, Stepansky et al. (1999) used inter-SSR-PCR and RAPD techniques to detect differences between North American and European *Cantalupensis* and *Inodorus* cultivars and exotic melon subsp. *agrestis* genotypes (e.g., Conomon, Dudaïm, and Momordica).

In our study, RAPD and SSR loci were used to assess variation among accessions in two horticulturally important melon groups (*Cantalupensis* and *Inodorus*) that are genetically distinguishable (Staub et al., 1997a; Stepansky et al., 1999). The array of elite *Cantalupensis* and *Inodorus* germplasm included provided genetic estimates of accessions having a relatively narrow genetic base. Molecular markers have been used to discriminate some of these market classes (Garcia et al., 1998; Mo-Suk et al., 1998; Stepansky et al., 1999). The marker loci in our study allowed for the separation of the germplasm examined into broad market class groupings (Figure 2). However, these groupings were somewhat ambiguous, agreeing with findings of Stepansky et al. (1999). The lack of complete partitioning of accessions of specific market classes into unique groupings likely reflects pedigree relationships and the introgression of specific traits between market classes during plant improvement.

It would be useful to know the composition (i.e., ancestral lines) of germplasm being examined during diversity analysis in order to allow for the introgression of unique variation (genes) from selected accessions (i.e., heirloom cultivars) into elite populations, and/or define reference accessions for future comparisons. 'Top Mark' and several Galia class accessions define the limits of the genetic diversity in the limited germplasm array examined in our study. A knowledge of their parentage would allow for greater understanding during breeding and future diversity assessment. 'Top Mark' was developed commercially by Robert Tang of Desert Seed Company, El Centro, Calif. and its pedigree is not known (personal communication, J.D. McCreight, USDA, ARS, Salinas, Calif.). The pedigree of 'Galia' is complex and represents a unique recently developed market class characterized by a round, yellow, finely netted fruit possessing a green flesh (Table 1). The original 'Galia' (F1) that was released in 1974 by Zvi Karchi [Newe Ya'ar Research

Center, Agricultural Research Organization (ARO), Israel], was developed from an Ogen type. Since then public institutions and private industry have developed proprietary cultivars of the 'Galia' type using various breeding strategies (e.g., Ogen  $\times$  Charentais) (personal communication, Zvi Karchi, ARO, Israel, 1998). 'Noy Yizre'el' (maternal) and an line of Russia origin are the parents of the original 'Galia' melon. 'Noy Yizre'el' originated from a cross between 'Ha'Ogen' possessing a green flesh color and 'Seminole', an U.S. eastern 'Rock Melon' type (developed for Florida conditions) having blue-green skin, light-moderate net, orange flesh, powdery mildew resistance and a small seed cavity. Backcrosses to 'Ha'Ogen' and several generations of inbreeding restored the Ogen type. The pedigree relationships between Galia and Ogen types in our study is supported by the small GD estimates presented herein, and suggests that the commercial lines/hybrids examined likely have Galia and Ogen market types in their pedigrees. The fact that some Galia and Ogen market types examined herein are distinct from other accessions within and among market classes of Group *Cantalupensis* and *Inodorus* suggests the introgression of genes (i.e., disease resistance, improved shipping and keeping quality) from more diverse germplasm in the germplasm examined.

Effective germplasm management depends upon a genetic understanding of relatedness and the frequency of genotypes in a population. Likewise, a knowledge of genetic relationships is important in plant variety protection (PVP) because intellectual ideas may be similar and infringement of proprietary rights occasionally occurs (Staub et al., 1996c). In both cases, GD estimation is important in the determination of distinctiveness. The terms absolute GD (a perfect reflection of reality as defined by the locus by locus evaluation of genomic differences), relative GD (an estimation of reality based on defined data sets), and functional GD (an estimation that accounts for a significant amount of the observed variation for a trait) are used to describe genetic relationships. Estimations of absolute GDs are cost prohibitive, and estimations of functional GD are of limited value unless a trait(s) is uniquely distinctive.

An extensive database must be developed for comparative analysis of germplasm in order to obtain more accurate estimates of GD in melon (Staub and Meglic, 1993). We present, herein, a comparative estimate of relative GDs in melon which together with those estimates of Garica et al. (1998) provide a foundation for further analysis of melon diversity. The

partitioning of accessions into major *C. melo* subsp. groupings (*Cantalupensis* and *Inodorus*) with RAPD markers was observed in this study and that of Staub et al. (1997a). Likewise, the unique position of Casaba accessions after hierarchical analysis in our study (Figure 2) supports that found by Garcia et al. (1998). The unique molecular character of Ogen and Shipping types was also revealed by the present study. Nevertheless, it is questionable whether such relative GD estimations provide adequate discrimination for use in PVP (Staub, 1999). These estimations must be viewed in comparison with more comprehensive molecular characterizations of intra- and intermarket class variation (i.e., other market classes such as African, Middle East, and Oriental types) as they become available.

Our study documents the use of bulked sampling for germplasm assessment using RAPD and SSR markers in melon. Individuals within that particular bulk can be sampled to estimate allelic frequencies where bulk sampling reveals the presence of multiple SSR alleles at a locus. This reductionism strategy maximizes the use of rapid bulk sampling to characterize the allelic nature of accessions for further analysis with cost minimization.

## Acknowledgements

This research is supported in part by a grant No. IS-2708-96 from the US-Israel Binational Agricultural Research and Development (BARD) Fund, and a grant from seven commercial vegetable seed companies.

## References

- Akagi, H., Y. Yokozeki, A. Inagaki and T. Fujimura, 1997. Highly polymorphic microsatellites of rice consist of AT repeats and a classification of closely related cultivars with these microsatellite loci. *Theor Appl Genet* 94: 61–67.
- Baudracco-Arnas, S. and M. Pitrat, 1996. A genetic map of melon (*Cucumis melo* L.) with RFLP, RAPD, isozyme, disease resistance and morphological markers. *Theor Appl Genet* 93: 57–64.
- Bailey, L.H. and E.Z. Bailey, 1976. *Hortus Third*, MacMillan Publishers, New York.
- Beckmann, J.S. and M. Soller, 1990. Marker-based mapping of quantitative trait loci using replicated progenies. *Theor Appl Genet* 80: 205–208.
- Bretting, P.K. and M.P. Widerlechner, 1995. Genetic markers and horticulturalgermplasm management. *HortScience* 30: 1349–1356.
- Davies K., 1993. Of mice and men (and cows and cats). *Nature* 361: 478.



- Debener, T., F. Salamini and C. Gebhart, 1990. Phylogeny of wild and cultivated *Solanum* species based on nuclear restriction fragment length polymorphisms (RFLPs). *Theor Appl Genet* 79: 360–368.
- Diwan, N. and P.B. Cregan, 1997. Automated sizing of fluorescent-labelled simple sequence repeat (SSR) markers to assay genetic variation in soybean. *Theor Appl Genet* 95: 723–733.
- Efron, B. and R. Tibshirani, 1986. Bootstrap methods for standard errors, confidence limits and other measures of statistical accuracy. *Stat Sci* 1: 54–77.
- Esquinas-Alcázar, J.T., 1977. Alloenzyme variation and relationships in the genus *Cucumis*. Ph.D. Diss., Univ California, Davis, 170 pp.
- García E., M. Jamilena, J.I. Alvarez, T. Arnedo, J.L. Oliver and R. Lozano, 1998. Genetic relationships among melon breeding lines revealed by DNA markers and agronomic traits. *Theor Appl Genet* 96: 878–885.
- Gibbons, J.D., 1976. Nonparametric Methods for Quantitative Analysis. Holt, Rinehart and Winston. New York. pp. 463.
- Hubalek, Z., 1982. Coefficients of association and similarity, based on binary (presence-absence) data: An evaluation. *Biol Rev Camb Philos Soc* 57: 669–689.
- Jackson D.A., K.M. Somers and H.H. Harvey, 1989. Similarity coefficients: Measurements of co-occurrence and association or simply measures of occurrence? *Am Nat* 133: 436–453.
- Janowitz, M.F., 1980. Similarity measures on binary data. *Syst Zool* 29: 342–359.
- Jeffery, C., 1980. A review of the Cucurbitaceae. *Bot J Linn Soc* 81: 233–247.
- Katzir N., Y. Danin-Poleg, G. Tzuri, Z. Karchi, U. Lavi and P.B. Cregan, 1996. Length polymorphism and homologies of microsatellites in several Cucurbitacea species. *Theor Appl Genet* 93: 1282–1290.
- Maniatis, T., E.F. Fritsch and J. Sambrook, 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor publisher, Cold Spring Harbor, NY.
- Mo-Suk, Y., H. Im-Sung, D. Go-Gawn, M. Ann-Chong, H. Kim-Doo, Y. Mo-Suk, and S.H. Im, 1999. RAPD analysis of genetic diversity of melon species. *Korean J Hort Sci Tech* 16: 21–24.
- Munger, H.M. and R.W. Robinson, 1991. Nomenclature of *Cucumis melo* L. *Cucurbit Genet Coop Rpt* 14: 43–44.
- Naudin, C.V., 1859. Essais d'une monographie des especes et des varieties du genre *Cucumis*. *Ann. Sci Natl Bot ser* 4, 11: 5–87.
- Nei, M., 1973. Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sci USA* 70: 3321–3323.
- Nei, M., 1987. Molecular evolutionary genetics. Colombia University Press, New York.
- Neuhausen, S.L., 1992. Evaluation of restriction fragment length polymorphisms in *Cucumis melo*. *Theor Appl Genet* 83: 379–384.
- Noli, E., S. Salvi and R. Tuberosa, 1997. Comparative analysis of genetic relationships in barley based on RFLP and RAPD markers. *Genome* 40: 607–616.
- Pitrat, M., 1994. Linkage groups in *Cucumis melo* L. *Cucurbit Genet Coop Rpt* 17: 148–149.
- Rafalski, J.A. and S.V. Tingey, 1993. Genetic diagnostics in plant breeding: RAPDs, microsatellites and machines. *Trends Genet* 9: 275–279.
- Rolf, F.J., 1997. NTSYS-pc: Numerical taxonomy and multivariate analysis system. Exeter Software, New York.
- Stepansky, A., I. Kovalski, R. Perl-Treves and C.V. Naudin, 1999. Intraspecific classification of melons (*Cucumis melo* L.) in view of their phenotypic and molecular variation. *Plt Syst Evol* 217: 313–332.
- Shattuck-Eidens, D.M., R.N. Bell, S.L. Neuhausen and T. Hellentjaris, 1990. DNA sequence variation within maize and melon: Observations from polymerase chain reaction amplification and direct sequencing. *Genetics* 126: 207–217.
- Sokal, R.R. and P.H. Sneath, 1963. Principles of Numerical Taxonomy. Freeman, San Francisco, California.
- Sokal, R.R. and F.J. Rohlf, 1981. Biometry. 2nd ed. W.H. Freeman, New York.
- Spooner, D.M., J. Tivang, J. Nienhuis, J.T. Miller, D.S. Douches and A. Contreras-M, 1996. Comparison of four molecular markers in measuring relationships among the wild potato relatives *Solanum* section *Etuberosum* (subgenus *Potato*). *Theor Appl Genet* 92: 532–540.
- Staub, J.E. and V. Meglic, 1993. Molecular genetic markers and their legal relevance for cultivar discrimination: A case study in cucumber. *HortTechnology* 3: 291–300.
- Staub, J.E., F.C. Serquen and M. Gupta, 1996a. Genetic markers, map construction and their application in plant breeding. *HortScience* 31: 729–741.
- Staub J.E., J. Bacher and K. Poetter, 1996b. Sources of potential errors in the application of random amplified polymorphic DNAs in cucumber. *HortScience* 31: 262–266.
- Staub, J.E., A. Gabert and T.C. Wehner, 1996c. Plant variety protection: A reconsideration of genetic relationships. *HortScience* 31: 1086–1091.
- Staub, J.E., J. Box, V. Meglic, T.F. Horejsi and J.D. McCreight, 1997a. Comparison of isozyme and random amplified polymorphic DNA data for determining intraspecific variation in *Cucumis*. *Genet Res Crop Evol* 44: 257–269.
- Staub, J.E., F.C. Serquen and J.D. McCreight, 1997b. Genetic diversity in cucumber (*Cucumis sativus* L.): III. An evaluation of Indian germplasm. *Genet Res Crop Evol* 44: 315–326.
- Staub, J.E., 1999. Intellectual property rights, genetic markers and hybrid seed production. *J New Seeds* 1: 39–64.
- Steel, R.G.D. and J.H. Torrie, 1980. Principles and Procedures of Statistics: Abiometrical Approach. Second edition. McGraw-Hill, Inc., New York.
- Tivang, J., J. Nienhuis and O.S. Simth, 1994. Sampling variances of molecular marker data sets using the bootstrap. *Theor Appl Genet* 89: 259–264.
- Waugh, R. and W. Powell, 1992. Using RAPD markers for crop improvement. *Trends in Biotech* 10: 186–191.
- Whitaker, T.W. and D.W. Davis, 1962. Cucurbits: Botany, Cultivation and Utilization. Interscience Pub., New York.
- Wilkinson, L., 1989. SYSTAT: The System for Statistics. SYSTAT Inc., Evanston, Illinois.
- Williams, J.G.K., A.R. Kubelik, K.J. Livak, J.A. Rafalski and S.V. Tingey, 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucl Acids Res* 18: 6531–6535.
- Wu, K.S. and S.D. Tanksley, 1993. Abundance, polymorphism and genetic mapping of microsatellites in rice. *Mol Gen Genet* 241: 225–235.
- Xiao, J., L. Li, L. Yaun and S.R. McCouch, 1996. Genetic diversity and its relationship to hybrid performance and heterosis in rice as revealed by PCR-based markers. *Theor Appl Genet* 92: 637–643.

